Structural Properties of Optimal Transmission Policies over a Randomly Varying Channel

Mukul Agarwal

Laboratory for Information and Decision Systems,

M.I.T., 77 Mass. Ave., Cambridge, MA 02139, USA

(magar@MIT.EDU).

Vivek S. Borkar, Fellow (IEEE),

School of Technology and Computer Science,

T.I.F.R., Homi Bhabha Road, Mumbai. 400005, India.

(borkar@tifr.res.in).

Abhay Karandikar, Member (IEEE)

Department of Electrical Engineering,

I.I.T. Bombay, Powai, Mumbai, 400076, India

(karandi@ee.iitb.ac.in).

Abstract

We consider the problem of transmitting packets over a randomly varying point to point channel with the objective of minimizing the expected power consumption subject to a constraint on the average packet delay. By casting it as a constrained Markov decision process in discrete time with time-averaged costs, we prove structural results about the dependence of the optimal policy on buffer occupancy, number of packet arrivals in the previous slot and the channel fading state for both i.i.d. and Markov arrivals and channel fading. The techniques we use to establish such results: convexity, stochastic dominance, decreasing-differences, are among the standard ones for the purpose. Our main contribution, however, is the passage to the average cost case, a notoriously difficult problem for which rather limited results are available. The novel proof techniques used here are likely to have utility in other stochastic control problems well beyond their immediate application considered here.

Index Terms

Randomly varying channel, transmission scheduling, power control, time averaged cost, constrained Markov decision processes.

I. INTRODUCTION AND RELATED WORK

Power efficient communication has been an important design challenge for wireless communications. While there are avenues for power savings in transmitters at a variety of implementation stages including radio circuitry, communication protocols etc, in this paper we concern ourselves with the power savings that can be achieved through packet scheduling where the transmitter gains by transmitting packets at a more opportune time or in a more opportune fashion.

In wireless as well as in wired channels with Additive White Gaussian Noise (AWGN), at the physical layer, the transmission power required for reliable communication increases as a convex function of the transmission rate. The convex nature of the relationship between power and rate allows one to save energy by choosing the rate at various stages in transmission in an appropriate fashion. This, however, would lead to an increase in buffered data and hence the average delay.

The wireless channel being time varying offers another opportunity of power savings where the scheduler can simply defer the transmission of packets during 'bad' channel states to 'good' channel states. In this paper, we consider the problem of scheduling packets over a point to point channel. The objective is to minimize the average power consumption subject to a constraint on the average packet delay. The problem of energy efficient scheduling for a wireless channel has been considered in [1], [2], [3], [4], [5], [6], [7], [8], [9], [10].

In [7], the authors have considered the problem of power efficient scheduling under average and absolute delay constraints. They have used a slotted system and considered an arbitrary independent and identically distributed (i.i.d.) packet arrival process. A characterization of the optimal scheduler has been provided in terms of a smaller class of deterministic schedulers. In [8] and [9], the authors consider energy minimization over a given time interval in a wireless network. Both off-line static optimization algorithms and on-line heuristics are considered. In [4], the authors have dealt with energy efficient scheduling under an average delay constraint for time-varying channels. However, the authors have assumed a linear relationship between power (P) and rate (R), and thus do not take into account the gains that can be realized by varying the transmission rates.

In [2], [3], the tradeoff between the average delay and the average power in a fading channel has been analyzed. The delay-power tradeoff has also been quantified in the region of asymptotically large delays. In [10] also, the author has considered the problem of energy efficient scheduling taking into account both the fluctuating channel conditions and the convex $P - R$ relationship. The structural results for a policy which minimizes the average delay subject to a constraint on the average power, in presence of channel fading, are provided in [5]. It is proved in [2] and [5] that there exists an optimal stationary policy which increases as the buffer occupancy increases, and decreases as the channel state goes from good to bad. What this means in physical terms is that the optimal decision is to transmit a certain number of packets (or, in our continuous model, the 'fluid' approximation thereof) at any given instant, where this number is an increasing function of the current queue length and a decreasing function of the channel state. Thus for a fixed channel gain, the greater the queue length the more you transmit, and for a fixed queue length, the better the channel, the more you transmit. This interpretation will be used throughout when we talk of 'increasing' and 'decreasing' policies. See also [1], [6] for some more results in this vein: these articles also derive explicit structural results for optimal policies under specific set of assumptions.

Our model is similar to that of [2] with a single transmitter and a single receiver on a point to point wireless link with fading. The only difference is that we assume fluid packet arrival and departure processes. We have also considered the arrival process to be both i.i.d. and first order Markov. The channel state has also been considered for both cases - i.i.d. and first order Markov. The primary contribution of the paper is to derive structural results of an optimal policy that minimizes the average power subject to an average delay constraint. Specifically, we improve upon the results of [2], [5]. The results not available in [2], [5] are concerning Markovian packet arrivals. They are:

   1) In case of First order Stochastically Dominant (FSD) Markov arrivals, the optimal policy is increas-

ing in the number of packet arrivals in the previous slot.

2) The existence of a stationary policy for the average cost problem when the packet arrival process is Markovian.

The existence of a stationary optimal policy for the average cost problem when the packet arrival process is i.i.d. has been proved in [5], but the problem becomes much more difficult when the arrival process is Markovian. This is because the state of the arrival process is an additional state variable on which the policy must depend, and the dynamics of this state variable may not be as explicit as that of the queue length process. Also, [5] does not derive the average cost dynamic programming equation which characterizes the optimal policy.

In a recent work [11], a discrete state space version of this problem is analyzed. Our problem may be considered a time-discretized version of the so called fluid limit of the discrete problem. However, in practice, channel fading state will be continuous valued (discretization is usually an approximation). Even [11] has mentioned this see, e.g., their Example 2. This automatically makes the state space non-discrete, as this is one component of the state space. Once the state space is continuous even in one component, the usual (easier) treatment of average cost dynamic programming equation does not apply. Denumerable or general state space problems involve in particular nontrivial stability considerations, involving, e.g., the choice of initial guess in policy iteration. See the work of Sean Meyn on this [12].

It is true that the packet arrivals are likely to be discrete in practice but the 'fluid' approximations of arrivals such as ours are commonplace. Our treatment can also cover the discrete arrivals / queue length case - the only part that does not go through is the uniqueness of minimizer in the dynamic programming equation, implying that the optimal policy may be randomized. We believe that fluid limits have the advantage that a simple fluid limit often serves as a robust approximation to a wide class of discrete state models. The reason for this is that the limit theorems through which the fluid limits are arrived at depend on gross local characteristics like the conditional mean, and therefore suppress other details that are irrelevant in the limit. This allows for the same fluid limit to work for a class of models rather than a single model. Furthermore, for the same reason, the fluid limit usually offers a more compact description.

Our approach will be to begin with the finite horizon discounted problem and pass first to the infinite horizon limit, followed by the vanishing discount limit to obtain the average cost case. There is copious literature on structural results for optimal policies in Markov decision processes, see, e.g., [13], [14], [15], [16] among others. Our main contribution is the passage from the infinite horizon discounted case to the average cost case, a notoriously difficult problem for which rather limited results are available. While the

average cost problem for general state spaces has been extensively studied in stochastic control literature [17], [18], the results available use conditions either too restrictive or not easily verifiable for problems such as the one studied in this paper. The novel proof techniques used here are likely to have utility in other stochastic control problems well beyond their immediate application considered here. Specifically, the following are the highlights of our proof technique:

1) We combine 'coupling at pseudo-atom' argument with a pathwise comparison argument based on stochastic dominance. While both these techniques exist separately (the former in fact is not so commonly known or used), this is the first time they are thus combined. The combination provides a more concise argument for the scalar case, which is also intuitively more appealing.

2) The passage to the vanishing discount limit for the general state space has to be based upon an equicontinuity-boundedness argument that invokes Arzela-Ascoli theorem (See p. 214 of [19]), as the simple argument based on Bolzano-Weierstrass theorem (See p. 77 of [19]) in discrete framework does not work. The argument given by us which is based on showing that the renormalized discounted value attains its minimum in a prescribed bounded set independent of the discount factor is new in discrete time framework and is applicable more generally to all cost criteria that penalize large excursions of the state process.

The rest of the paper is organized as follows. Section II discusses the model, problem formulation and the solution approach. We summarize our main contributions in Section III. The subsequent sections provide the proofs of the main results. Specifically, the unconstrained finite horizon discounted cost problem and the unconstrained infinite horizon discounted cost problem are considered in Sections IV and V. The unconstrained average cost problem is then addressed in Section VI. The paper concludes in Section VII.

II. Model, Problem Formulation and Approach

A. System Model

There is a point to point channel over which packets are being transmitted. Packets arrive at the transmitter with a queue of infinite size and get buffered. The system is discrete time, the dynamics of which is given by:

$$x_{n+1} = x_n - y_n + w_{n+1}.$$

Here x_n is the queue length or buffer occupancy at the beginning of slot n, y_n is the number of packets transmitted in this slot, and w_{n+1} the number of new arrivals in this slot taking values in a finite interval $[0, w_{max}]$ where w_{max} is the maximum number of arrivals in a slot. x_n, y_n and w_n are assumed to be fluid, i.e., continuous valued. The channel is time varying with fading. The state of the channel indicates the time varying channel gain. We assume flat fading model, i.e., the channel state is assumed to remain constant over the slot duration (See [2] for a detailed discussion on this). The packet arrival process and channel state evolution process are assumed to be stationary and independent. Let $\{\mu_n\}$ denote the channel state process taking values in $\mathbb{R}$. We shall consider both i.i.d. and Markovian arrivals with law (resp., transition kernel) $p(dw)$ (resp., $p(a, dw)$) and both i.i.d. and Markovian channel with law (resp., transition kernel) $q(d\mu)$ (resp., $q(\nu, d\mu)$). The energy required to transmit y packets under channel state μ will be assumed to be $\mu F(y)$ for a convex and increasing $F : \mathbb{R}^+ \to \mathbb{R}^+$. Let w_0 denote the stationary average of the number of packet arrivals in a slot. Let y_{max} denote the maximum number of packets that can be transmitted in a slot. $w_{max}, w_0, y_{max} \in \mathbb{R}$. For the stability of the buffer, $y_{max} > w_0$.

The state of the above system can be completely characterized by the 3-tuple, $\mathbf{v_n} = (x_n, w_n, \mu_n)$, comprising of the buffer occupancy or queue length, the number of packet arrivals in the previous time slot and the channel state. In slot n, the control (or scheduling) action corresponds to the number of packets transmitted, y_n. The control policy is a sequence of functions $\{\pi_1, \pi_2, \ldots\}$ where π_n specifies the conditional law of y_n given the past history of the system state and the past applied controls, i.e., given $\mathbf{v_0}, \mathbf{v_1}, \ldots, \mathbf{v_n}, y_1, y_2, \ldots, y_{n-1}$.

B. Problem Formulation

Since the packets arrive and get queued in the buffer, they suffer a delay. By Little's theorem (Chapter 3 of [20]), the average packet delay, D, is related to the time-averaged queue length, Q, as:

$$D = \frac{Q}{w_0}, \tag{1}$$

where w_0 denotes the average packet arrival rate. Hence, in the rest of the paper, we ignore the proportionality constant w_0, and treat average delay as synonymous with average queue length. In our problem, Q can be written as:

$$Q = \limsup_{M \to \infty} \frac{1}{M} \Big[\sum_{n=1}^{M} x_n \Big]. \tag{2}$$

We thus define:

Definition 1—Average Delay, D; Average Power, P:

$$D = \limsup_{M \to \infty} \frac{1}{M} E\Big[\sum_{n=1}^{M} x_n \Big], \ \ P = \limsup_{M \to \infty} \frac{1}{M} E\Big[\sum_{n=1}^{M} \mu_n F(y_n) \Big] \tag{3}$$

Recall that the energy required to transmit packets is a convex function of the number of packets being transmitted (transmission rate). Thus, from an energy efficiency point of view, we would want to transmit packets in small chunks. Transmitting packets in small chunks leads to higher delay. Thus, we have the average cost optimal scheduling problem that can be stated as: Determine a scheduling policy that minimizes P subject to a constraint on D. It is a constrained Markov decision problem. To solve it, we first consider the corresponding unconstrained Markov decision problem.

Let λ be a positive real number. Let $C_\lambda = P + \lambda D$, where, P and D are defined in (3). The unconstrained average cost optimal scheduling problem can be stated as: Determine a scheduling policy that minimizes C_λ. Later on, we shall map the constrained problem to this unconstrained one by interpreting λ as the appropriate Lagrange multiplier.

Before we can solve this problem, we first address the corresponding infinite horizon discounted cost problem. Let $0 < \beta < 1$ be a real number. Define:

Definition 2—Infinite Horizon Expected Discounted Delay:

$$D_\infty = E\Big[\sum_{n=1}^{\infty} \beta^n x_n \Big] \tag{4}$$

Definition 3—Infinite Horizon Expected Discounted Power:

$$P_\infty = E\Big[\sum_{n=1}^{\infty} \beta^n \mu_n F(y_n) \Big] \tag{5}$$

Define $C_{\infty,\lambda} = P_\infty + \lambda D_\infty$. The infinite horizon unconstrained β-discounted cost optimal scheduling problem can be stated as: Determine a scheduling policy that minimizes $C_{\infty,\lambda}$. This problem is addressed by first formulating the corresponding finite horizon problem. We then address the unconstrained average cost and finally the constrained average cost problem.

C. Mathematical Preliminaries and Assumptions

In this section, we review some mathematical preliminaries. We first define the notion of *First Order Stochastic Dominance* (FSD). For probability measures p_1, p_2 on $\mathbb{R}$, p_1 is said to be First Order Stochastically Dominant with respect to p_2 if $p_1((-\infty, x)) \leq p_2((-\infty, x))$, $\forall x \in \mathbb{R}$. First Order Stochastic Dominance of p_1 with respect to p_2 will be denoted by $p_1 \succ p_2$, or by $p_2 \prec p_1$. The following two results are proved in Chapter 1 of [21].

Lemma 1: Let $p_1 \succ p_2$ be probability measures on $\mathbb{R}$. Let $f : \mathbb{R} \to \mathbb{R}$ be an increasing function. Then

$$\int_{\mathbb{R}} f(w) p_1(dw) \geq \int_{\mathbb{R}} f(w) p_2(dw).$$

Lemma 2: Let $p_1 \succ p_2$ be probability measures on $\mathbb{R}$. Then there exist random variables $\bar{X}_1$ and $\bar{X}_2$ on a common probability space such that $\bar{X}_1 \sim p_1$, $\bar{X}_2 \sim p_2$, and $\bar{X}_1 \geq \bar{X}_2$ almost surely.

We make the following assumptions for the incoming packet arrival distribution $p(a, dw)$.

1) $p(b, dw) \prec p(a, dw)$ whenever $a > b$.

2) $p(x, dw)$ is continuous in x in total variation norm: $\int |p(x, dw) - p(x', dw)| \overset{x' \to x}{\to} 0$.

The following more stringent conditions will be used in our analysis of the average cost problem:

1) $p(\cdot, \cdot)$ satisfies:

$$\int |p(a', dw) - p(a, dw)| \leq c_0 |a' - a|, \quad \text{for some constant } c_0. \tag{6}$$

2) The process $< X_n, W_n, \mu_n >$ is ergodic under any stable stationary policy.

3) Minorization Condition: Let $W = [0, w_{max}]$. There exists a probability measure $\hat{\nu}$ on W and a real number $\delta > 0$ such that for all $x \in W$, $A \subset W$ Borel,

$$p(x, A) - \delta \hat{\nu}(A) \geq 0. \tag{7}$$

Note 1: The minorization condition holds, for example, if there is a neighborhood N of 0 such that $p(x, N) \geq \epsilon$ for all x, where $\epsilon > 0$ is a constant. (Take $\hat{\nu} = $ the normalized uniform distribution on N.) Intuitively, this means that there is a strictly positive probability of "close to" zero packet arrivals irrespective of the number of packet arrivals in the previous slot.

III. SUMMARY OF MAIN RESULTS AND PROOF OUTLINE

We first summarize our main results and give a sketch of the methodology adopted for the proofs. The details of the proofs are given in the subsequent sections. Our main results are for the infinite horizon average cost problems, both constrained and unconstrained. We first consider i.i.d. channel fading.

A. Channel Fading- i.i.d.

Theorem 1: 1) The average cost dynamic programming equation for the unconstrained problem:

$$V(x, a, \mu) = \min_{y \leq x \wedge y_{max}} \left(\lambda x + \mu F(y) - \rho + \int q(d\nu) \int p(a, dw) V(x - y + w, w, \nu) \right), \tag{8}$$

has a solution $(V(\cdot,\cdot,\cdot),\rho)$, where ρ is uniquely characterized as the optimal cost and V is unique up to an additive constant on the support of the stationary law under any optimal stationary policy.

2) V is convex increasing in the buffer occupancy or queue length, increasing in the number of arrivals, and decreasing in the channel state. Also, it is supermodular in the first two arguments, i.e., queue length and the number of arrivals.

3) $y^*(x,a,\mu)$ given by

$$y^*(x,a,\mu) = \arg\min_{[0,x\wedge y_{max}]}\{\mu(F(y) + \int q(d\nu)\int p(a,dw)V(x-y+w,w,\nu)\},$$

is an optimal stable stationary strategy and is increasing in the queue length and the number of arrivals and decreasing in the channel state.

The proof strategy followed is:

1) Derive the dynamic programming equation for the finite horizon discounted cost control problem and establish the aforementioned monotonicity and supermodularity properties for the value function (Section IV).

2) Consider the infinite time limit of the above and justify the dynamic programming equation for the infinite horizon discounted cost problem. This involves verifying boundedness and equicontinuity of the finite horizon discounted value functions. The monotonicity and supermodularity properties for the infinite horizon discounted value function follow from the corresponding properties for the finite horizon problem.

3) Consider the vanishing discount limit of the dynamic programming equation for the infinite horizon discounted cost problem after suitable renormalization, which amounts to subtracting from the value function its value at a prescribed state. This yields the dynamic programming equation for the average cost problem. The proof involves boundedness and equicontinuity of the renormalized infinite horizon discounted value functions. This is done using a coupling argument based on the Athreya-Ney-Nummelin pseudo atom construction (Chapter 5 of [22]), combined with a pathwise comparison that uses stochastic monotonicity (the latter distinguishes it from the few earlier uses of coupling at pseudo atom in literature for similar purposes), and finally, a novel proof technique to prove that the renormalized infinite horizon discounted value functions and therefore the average cost value function are uniformly bounded from below, which in turn allows us to prove the first part of the theorem above (Section VI).

Note 2: Note that the use of Little's theorem to justify the particular cost function above is not valid for the finite and infinite horizon discounted cost problems. These, however, are merely intermediate steps to

our average cost problem for which it is indeed justified.

Our main concern, however, is the constrained problem for which we have the following result. Assume that F is strictly convex.

Theorem 2: There exists a unique stable stationary optimal policy $y^*(x, a, \mu)$ increasing in buffer occupancy, x and the number of arrivals, a and decreasing in the channel state, μ.

By a standard 'Lagrange multiplier' formulation (see, e.g., [23]), it follows that the constrained problem has a stationary, though possibly randomized, optimal policy which is also optimal for the unconstrained problem considered in Theorem 1 for a particular choice of $\lambda = \lambda^*$ (say), the Lagrange multiplier for the problem. But the optimal stationary policies for the latter, randomized or not, must attain the minimum on the right hand side (r.h.s.) of the Bellman equation (8) for each (x, a, μ). By the strict convexity of F, this minimum is attained at a unique point, whence there is a unique optimal stationary policy with the stated properties, viz., increasing in x, a, and decreasing in μ.

Note 3: While we do not consider computational issues here, it is worth making a few comments regarding these. The preferred computational technique for constrained Markov decision problems in the past has been the linear programming approach [24]. As the state space here is not discrete, it becomes an abstract infinite dimensional linear program, which needs an approximation step to reduce it to a finite linear program. While discretization methods have been proposed for this purpose in the past [25], the more recent approach based on function approximation [26] holds great promise. An alternative is to use 'primal-dual' type methods which use a conventional iteration scheme for the value function and a dual ascent for the Lagrange multiplier [27]. (It may be recalled here that the standard iterative schemes for Markov decision processes have been extended to general state spaces, see, e.g., [12].) These schemes can also be combined with a function approximation scheme for dimensionality reduction. In either, structural results such as the ones presented here aid greatly in the choice of basis functions ('features' in Artificial Intelligence parlance) in function approximation. Function approximation based approximate linear and dynamic programming is currently an active area of research. What is more, a typical situation is the one where the model is not known and an on-line learning scheme is warranted. Such schemes based on function approximation are no harder for continuous state space than for discrete state space.

B. Channel Fading- Markov

In case of Markov channel fading, by using exactly the same methods used for proving Theorem 1 and Theorem 2, it can be proved that there exists an optimal stationary policy which is increasing in buffer occupancy and the number of packet arrivals in the previous time slot. However, unlike the case of i.i.d.

channel fading where the policy is decreasing in the channel fading state, nothing can be said in general when the channel fading is Markovian. We give an intuitive reason for this:

Let $q(\nu_1, d\mu) \succ q(\nu_2, d\mu)$ if $\nu_1 > \nu_2$. Intuitively, this means that if the channel fading in the current slot is high, it is expected that the channel fading in the next slot will be higher than what it would be if the channel fading in the current slot is low. Thus, even if the channel fading in the current slot is high, we might be better off transmitting a larger number of packets in anticipation of a very bad channel in the future, compared to when the channel fading in the current slot is low. For this reason, the number of transmissions in a bad channel state might be higher compared to the number of transmissions in a good channel state.

In the rest of the paper, we provide the proofs for Theorems 1 and 2 by first considering the unconstrained finite horizon discounted cost, infinite horizon discounted cost and finally the average cost problem.

IV. UNCONSTRAINED FINITE HORIZON DISCOUNTED COST PROBLEM

In this section, we consider the finite horizon discounted cost version of our problem as a first step to prove our main results as outlined in the previous section. In this section, we consider the channel fading to be i.i.d. We prove some properties like continuity and convexity of the optimal finite horizon β-discounted cost function. We also prove that this function is increasing in queue length, number of packet arrivals in the previous slot, and decreasing in channel state.

Let $V_n^\beta(\mathbf{v})$ denote the optimal β-discounted n-step cost when the initial state is $\mathbf{v}$, defined as the infimum over all possible admissible controls. Let $J_{n,\pi}^\beta(\mathbf{v})$ denote the β-discounted n-step cost when the initial state is $\mathbf{v}$ and the policy used is π. In the rest of this section, β will be a constant, and we do not explicitly show the superscript β. Thus, $V_n^\beta(\mathbf{v})$ will be denoted as $V_n(\mathbf{v})$ and $J_{n,\pi}^\beta(\mathbf{v})$ will be denoted as $J_{n,\pi}(\mathbf{v})$.

$V_n, n \geq 0$, the finite horizon value function, is given by the Bellman equation

$$V_n(x, a, \mu) = \min_{y \leq x \wedge y_{max}} [\lambda x + \mu F(y) + \beta \int q(d\nu) p(a, dw) V_{n-1}(x - y + w, w, \nu)], \; n > 0, \quad (9)$$

with $V_0(x, a, \mu) = \lambda x$. Also define

$$h_n(y, x, a, \mu) = \lambda x + \mu F(y) + \beta \int q(d\nu) p(a, dw) V_{n-1}(x - y + w, w, \nu), \quad \forall n$$

Theorem 3: $V_n(x, a, \mu)$ for each n is convex continuous and increasing in x, and satisfies

$$|V_n(x', a, \mu) - V_n(x, a, \mu)| \leq \frac{\lambda}{1 - \beta} |x' - x|.$$

Proof: See Appendix -A for the proof. ∎

Theorem 4: $V_n(x, a, \mu)$ for each n is continuous increasing in a and satisfies

$$V_n(x, a', \mu) - V_n(x, a, \mu) \leq \frac{w_{max}\lambda\beta}{(1 - \beta)^2} \int |p(a', dy) - p(a, dy)|.$$

Proof: See Appendix -B for the proof. ∎

Note that y_{max} is the maximum number of packets that can be transmitted in a slot. Then we can easily prove the following theorem.

Theorem 5: $V_n(x, a, \mu)$ for each n is continuous increasing in μ and satisfies

$$|V_n(x, a, \mu) - V_n(x, a, \mu')| \leq |\mu - \mu'|F(\min(x, y_{max})).$$

We now state the following 'supermodularity' property.

Theorem 6: Let Δx and Δa be arbitrary non-negative real numbers. Then, for each n, V_n satisfies,

$$V_n(x + \Delta x, a + \Delta a, \mu) + V_n(x, a, \mu) \geq V_n(x, a + \Delta a, \mu) + V_n(x + \Delta x, a, \mu) \tag{10}$$

Proof: The proof is outlined in Appendix -C ∎

V. UNCONSTRAINED INFINITE HORIZON DISCOUNTED COST PROBLEM

In this section, we analyze the unconstrained infinite horizon discounted cost problem. Let $0 < \beta < 1$ be a real number and $V^\beta(x, a, \mu)$ denote the optimal infinite horizon β-discounted cost when the initial state is (x, a, μ).

Lemma 3: There exists a function $M(x, a, \mu)$ such that $V_n^\beta(x, a, \mu) \leq M(x, a, \mu)$ for all n.

Proof: Let π_0 be the policy which transmits nothing in any time slot, irrespective of the state. Then

$$V_n^\beta(x, a, \mu) \leq J_{n,\pi_0}^\beta(x, a, \mu) \leq J_{\pi_0}^\beta(x, a, \mu) = \sum_{i=0}^{\infty} \beta^i(\lambda x + iw_{max}) \triangleq M(x, a, \mu) < \infty.$$

∎

Note that V_n^β increases as n increases. Suppose $V_n^\beta \uparrow W^\beta$. From the corresponding estimates for the V_n's, we have

$$|W^\beta(x', a', \mu') - W^\beta(x, a, \mu)| \leq |\mu' - \mu|F(\min(y_{max}, x'))$$
$$+ \frac{w_{max}\beta\lambda}{(1 - \beta)^2} \int |p(a', dy) - p(a, dy)| + \frac{\lambda}{1 - \beta}|x' - x| \tag{11}$$

In particular, W^β is continuous. It will clearly be increasing in its arguments and convex in x for fixed a, μ. Also, by Dini's theorem [28], $V_n^\beta \uparrow W^\beta$ uniformly on compact sets. Consider the Bellman equation,

$$V_n^\beta(x, a, \mu) = \min_{y \leq x \wedge y_{max}} [\lambda x + \mu F(y) + \beta \int q(d\nu) p(a, dw) V_{n-1}^\beta(x - y + w, w, \nu)] \quad (12)$$
$$\triangleq \min_{y \leq x \wedge y_{max}} h_n^\beta(y, x, a, \mu)$$

Define:

$$h^\beta(y, x, a, \mu) \triangleq \lambda x + \mu F(y) + \beta \int q(d\nu) p(a, dw) W^\beta(x - y + w, w, \nu) \quad (13)$$

Let $n \to \infty$ on both sides in (12). By the monotone convergence theorem (p. 377 of [19]),

$$\lim_{n \to \infty} \int p(a, dw) q(d\nu) V_{n-1}^\beta(x - y + w, w, \nu) = \int p(a, dw) q(d\nu) W^\beta(x - y + w, w, \nu)$$

Since $V_{n-1}^\beta \to W^\beta$ uniformly on compact sets, $h_n^\beta(y, x, a, \mu) \to h^\beta(y, x, a, \mu)$ uniformly on compact sets. Hence

$$\lim_{n \to \infty} \min_{y \leq x \wedge y_{max}} h_n^\beta(y, x, a, \mu) = \min_{y \leq x \wedge y_{max}} h^\beta(y, x, a, \mu)$$

Thus W^β satisfies:

$$W^\beta(x, a, \mu) = \min_{y \leq x \wedge y_{max}} [\lambda x + \mu F(y) + \beta \int q(d\nu) p(a, dw) W^\beta(x - y + w, w, \nu)] \quad (14)$$

(14) is the Bellman equation for the Infinite Horizon Discounted Cost Problem. By standard arguments, we have $V^\beta = W^\beta = \lim_{n \to \infty} V_n^\beta$, and,

$$V^\beta(x, a, \mu) = \min_{y \leq x \wedge y_{max}} [\lambda x + \mu F(y) + \beta \int q(d\nu) p(a, dw) V^\beta(x - y + w, w, \nu)] \quad (15)$$

Note also that since the infinite horizon Bellman equation is satisfied, there exists an optimal *stationary* policy for the infinite horizon β-discounted problem, given by

$$y^*(x, a, \mu) = \arg \min_{y \leq x \wedge y_{max}} [\lambda x + \mu F(y) + \beta \int q(d\nu) p(a, dw) V^\beta(x - y + w, w, \nu)] \quad (16)$$

Lemma 4: For all $\mu, x' \geq x, a' \geq a$,

$$V^\beta(x', a', \mu) + V^\beta(x, a, \mu) \geq V^\beta(x, a', \mu) + V^\beta(x', a, \mu). \quad (17)$$

This follows from the corresponding property of V_n^β from Theorem 4 by passing to the $n \to \infty$ limit therein.

VI. UNCONSTRAINED AVERAGE COST PROBLEM

In this section, we prove the existence of an optimal stable stationary policy for the unconstrained average cost problem. We do so under the additional assumptions of Section III.

The proof technique considers a suitably renormalized infinite horizon discounted value function. We first prove the boundedness and equicontinuity of this renormalized value function. This is proved in Lemma 12. To prove this lemma, we develop the arguments in Lemma 5, 6, 7, 8, 9, 10 and 11. This follows Athreya-Ney-Nummelin pseudo-atom construction (Chapter 5 of [22]) combined with a pathwise comparison. This is different from the earlier uses of pure 'coupling at pseudo-atom' argument. We then use a novel proof technique to show that the renormalized infinite horizon value functions and thereby the average cost value functions are bounded from below. This is proved in subsequent Lemma 13 and Theorem 8 and 38. We can thus prove our main results, i.e., Theorems 1 and 2. We develop these arguments in this section.

Recall that a stationary policy is stable if the one dimensional marginals of the corresponding Markov process remain tight. If so, a stationary distribution exists. The process is ergodic if this stationary distribution is unique. Then the time averages of functions of the Markov process that are integrable w.r.t. the stationary distribution converge a.s. to their respective averages w.r.t. this distribution. Under our minorization condition, stationarity automatically implies ergodicity.

Theorem 7: There exists a stationary distribution π for the Markov Chain $< W_n >$. Also, there exists $c > 0$, $0 < \gamma < 1$ such that

$$| \int w[p^{(n)}(a, dw) - \pi(dw)]| \leq c\gamma^n, \quad \forall n > 0. \tag{18}$$

A proof can be found in [22]. The additional burden of verifying 'geometric ergodicity' therein is unnecessary here because we have a compact state space anyway. Next, we state a result which will be used to prove the subsequent lemma.

Lemma 5: For all a, μ,

$$|V^\beta(x', a, \mu) - V^\beta(x, a, \mu)| = \mathrm{O}[|x' - x| + |x' - x|(x' + x)]. \tag{19}$$

Proof: Let $x' > x$ and $(x' - x) \leq \frac{y_{max} - w_0}{2} \triangleq \Delta$ where w_0 denotes the stationary expectation of the arrival sequence $< W_k >$ and y_{max} denotes the maximum number of packets that can be transmitted

in a slot. W_k is the arrival process with $W_0 = a$. Let $< Y_k >$ be optimal for the process:

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, W_0 = a, \mu_0 = \mu.$$

Let τ be the index of the first slot such that $Y_n \leq (y_{max} - \Delta)$ (recall that the slots are numbered starting with 0). Consider the process $< X_k' >$ defined inductively as:

$$X_{k+1}' = X_k' - Y_k' + W_{k+1}, X_0 = x'$$

where

$$Y_k' = \begin{cases} x' - x + Y_k & \text{if } \tau = k, \\ Y_k & \text{otherwise.} \end{cases}$$

The channel state process $< \mu_n >$ remains the same for both. It is easy to check that for $n \leq \tau$, $(X_n' - X_n) = (x' - x)$, and for $n > \tau$, $(X_n' - X_n) = 0$. Using the fact that $< Y_k' >$ need not be optimal for $< X_k' >$, and that, $V^\beta(x', a, \mu) \geq V^\beta(x, a, \mu)$, it is easy to see that

$$V^\beta(x', a, \mu) - V^\beta(x, a, \mu) \leq E[(\tau + 1)\lambda(x' - x)] + \mu_{max}[F(x' - x + Y_\tau) - F(Y_\tau)]$$

$$\leq \lambda(x' - x)(E\tau + 1) + \eta' \mu_{max}(x' - x) = \lambda(x' - x)E\tau + \eta(x' - x), \quad (20)$$

where η' is the Lipschitz constant for $F(y)$, $0 \leq y \leq y_{max}$ and $\eta = (\mu_{max}\eta' + \lambda)$. Here, μ_{max} is the maximum possible channel fading in a slot.

We now derive an upper bound on $E[\tau]$; more precisely, we prove that $E[\tau] \leq C_1 + C_2 x$ for suitable constants $C_1, C_2 > 0$. Note that $X_n \leq (y_{max} - \Delta) \Rightarrow Y_n \leq (y_{max} - \Delta) \Rightarrow \tau \leq n$. Let $\hat{\tau} =$ the index of the first slot when $X_n \leq y_{max} - \Delta$. Note that $X_m \geq (y_{max} - \Delta)$ for $m < \tau$, hence $\tau \leq \hat{\tau}$, a fact we shall use in order to derive an upper bound on $E[\tau]$.

Let $n_0 > 1$ be such that for $n \geq n_0$,

$$E(W_1 + W_2 + \cdots + W_n | W_0 = w) \leq n\left(w_0 + \frac{\Delta}{2}\right) \quad (21)$$

This is possible because $\frac{1}{n} E[\sum_{m=1}^n W_m | W_0 = w] \to w_0$.

Let $Z_n = X_{nn_0}$. Let $Z_n \geq n_0 y_{max}$ and $\hat{\tau} > nn_0$. Note that if $Z_n \geq n_0 y_{max}$ and $\tau > nn_0$, $\tau \geq (n+1)n_0$. This is because the maximum number of packets that can be transmitted in a slot is y_{max}

and therefore $X_m \geq y_{max}$ for $nn_0 \leq m < (n+1)n_0$.

$$E(Z_{n+1} \mid Z_n) - Z_n$$
$$= E(-Y_{nn_0} + W_{nn_0+1} - Y_{nn_0+1} + W_{nn_0+2} - \cdots - Y_{(n+1)n_0-1} + W_{(n+1)n_0} \mid Z_n)$$
$$= E(W_{nn_0+1} + \cdots + W_{(n+1)n_0} \mid Z_n) - E(Y_{nn_0} + \cdots + Y_{(n+1)n_0-1} \mid Z_n)$$
$$\leq^* n_0 \left(w_0 + \frac{\Delta}{2} \right) - n_0(y_{max} - \Delta)$$
$$= -n_0 \frac{\Delta}{2}.$$

(*) follows by (21) and the fact that $Y_n > (y_{max} - \Delta)$ for $n < \hat{\tau}$. Thus, $V_{liap}(x) = x$ is a Liapunov function for the process $< X_k >$ for $k < \hat{\tau}$. A standard argument (See p. 268 in [22]) then shows that the expected hitting time of the set $[0, n_0 y_{max}]$ for the process $< X_k >$ is bounded by $K(1 + x)$ for some $K > 0$ for all x. Next, we prove that if $x \leq n_0 y_{max}$, there exists a constant C such that $E[\hat{\tau}] \leq C$. Combining the two, it will follow that $E[\hat{\tau}]$ and hence $E[\tau]$ is bounded by $K'(1 + x)$ for some $K' > 0$ for arbitrary x. Thus we assume that $x \leq n_0 y_{max}$. Assume also that $x > (y_{max} - \Delta)$ since otherwise $\hat{\tau} = 0$. For $n \geq n_0$,

$$\Pr[\tau > n | X_0 = x, W_0 = w]$$
$$=^{*1} \Pr[X_0, X_1, \ldots, X_n > (y_{max} - \Delta) | X_0 = x, W_0 = w]$$
$$= \Pr[x > (y_{max} - \Delta), (x - Y_0 + W_1) > (y_{max} - \Delta), \ldots,$$
$$(x - Y_0 + W_1 - Y_1 + W_2 - \ldots - Y_{n-1} + W_n) > (y_{max} - \Delta) | W_0 = w]$$
$$\leq^{*2} \Pr\left[x + \sum_{i=1}^{n} W_i - n(y_{max} - \Delta) > (y_{max} - \Delta) \Big| W_0 = w \right]$$
$$= \Pr\left[\sum_{i=1}^{n} W_i - E \sum_{i=1}^{n} W_i > (y_{max} - \Delta) + n(y_{max} - \Delta) - x - E \sum_{i=1}^{n} W_i \Big| W_0 = w \right]$$
$$\leq^{*3} \Pr\left[\sum_{i=1}^{n} W_i - E \sum_{i=1}^{n} W_i > (y_{max} - \Delta) + n(y_{max} - \Delta) \right.$$
$$\left. -x - n\left(w_0 + \frac{\Delta}{2} \right) \Big| W_0 = w \right]$$
$$= \Pr\left[\sum_{i=1}^{n} W_i - E \sum_{i=1}^{n} W_i > n\left(\frac{\Delta}{2} + \frac{y_{max} - \Delta - x}{n} \right) \Big| W_0 = w \right].$$

- $(*^1)$ follows by the fact that $\hat{\tau} > n \implies X_m \geq (y_{max} - \Delta)$.
- $(*^2)$ follows since for $n < \tau$, $Y_n \geq (y_{max} - \Delta)$.

- $(*^3)$ follows by (21).

Let n_1 be such that $n_1 \geq n_0$, $\frac{y_{max} - \Delta - n_0 y_{max}}{n} \geq -\frac{\Delta}{4}$ for $n \geq n_1$, and $n \geq \frac{2w_{max}}{\delta\left(\frac{\Delta}{2} + \frac{y_{max} - \Delta - n_0 y_{max}}{n}\right)}$ for $n \geq n_1$. By the Hoeffding's inequality for Markov chains [29], for $n \geq n_1$,

$$\Pr[\hat{\tau} > n | X_0 = x, W_0 = w] \leq \exp\left[\frac{-\delta^2 \left(n\frac{\Delta}{4} - \frac{2w_{max}}{\delta}\right)^2}{2nw_{max}^2}\right]. \tag{22}$$

Thus, for $x \leq n_0 y_{max}$,

$$
\begin{aligned}
E[\hat{\tau} | X_0 = x, W_0 = w] &= 1 + \sum_{i=1}^{\infty} \Pr[\tau > i | X_0 = x, W_0 = w] \\
&\leq n_1 + \sum_{i=n_1}^{\infty} \exp\left[-\frac{\delta^2 \left(i\frac{\Delta}{4} - \frac{2w_{max}}{\delta}\right)^2}{2iw_{max}^2}\right] \\
&= C
\end{aligned}
$$

for some $C < \infty$. This is because the ith term of the series on the right is $O(e^{-Ki})$ for a suitable $K > 0$ independent of x, w and the series is therefore summable uniformly w.r.t. these variables. Thus for $x \leq n_0 y_{max}$, $E[\hat{\tau}] \leq C$. We had proved earlier that for arbitrary x, the hitting time of the set $[0, n_0 y_{max}]$ for the process $< X_k >$ is $O(x)$. It follows that $E[\tau] = O(x)$ for arbitrary x whenever $(x' - x) \leq \Delta$. We write this fact as: $E[\tau] \leq \alpha x + \eta$ for some suitable constants $\alpha, \eta > 0$.

We now remove the assumption that $(x' - x) \leq \Delta$.

$$
\begin{aligned}
&V^\beta \left(x', a, \mu\right) - V^\beta \left(x, a, \mu\right) \\
&= \sum_{n=0}^{\lfloor \frac{x' - x}{\Delta} \rfloor - 1} \left[V^\beta \left(x + (n+1)\Delta, a, \mu\right) - V^\beta \left(x + n\Delta, a, \mu\right)\right] \\
&\qquad\qquad + \left[V^\beta \left(x', a, \mu\right) - V^\beta \left(x + \left\lfloor \frac{x' - x}{\Delta} \right\rfloor \Delta, a, \mu\right)\right] \\
&\leq \sum_{n=0}^{\lfloor \frac{x' - x}{\Delta} \rfloor - 1} \left[\lambda\alpha \left(x + n\Delta\right) + \eta\right]\Delta + \left(\lambda\alpha x' + \eta\right)\left(x' - x - \left\lfloor \frac{x' - x}{\Delta} \right\rfloor \Delta\right) \\
&= \left(\lambda\alpha x + \eta\right)\Delta \left\lfloor \frac{x' - x}{\Delta} \right\rfloor + \frac{\lambda\alpha\Delta^2}{2}\left(\left\lfloor \frac{x' - x}{\Delta} \right\rfloor - 1\right)\left(\left\lfloor \frac{x' - x}{\Delta} \right\rfloor\right) \\
&\qquad + \left(\lambda\alpha x' + \eta\right)\left(x' - x - \left\lfloor \frac{x' - x}{\Delta} \right\rfloor \Delta\right) \\
&\leq \left(\lambda\alpha x + \eta\right)\left(x' - x\right) + \frac{\lambda\alpha}{2}\left(x' - x\right)^2 + \left(\lambda\alpha x' + \eta\right)\left(x' - x\right) \\
&= 2\eta\left(x' - x\right) + \lambda\alpha\left(x' + x\right)\left(x' - x\right) + \frac{\lambda\alpha}{2}\left(x' - x\right)^2 \\
&= O\left[|x' - x| + |x' - x|(x' + x)\right].
\end{aligned}
$$

This proves the lemma. ∎

Definition 4: $\zeta^{(n)} = \int w|p^{(n)}(a', dw) - p^{(n)}(a, dw)|.$

Lemma 6:   i) $\zeta^{(n)} \le k_1|a' - a|,$

ii) $\zeta^{(n)} \le k_2\gamma^n,$

for some constants k_1 and k_2, and γ is as in Theorem 7.

Proof: i)

$$
\begin{aligned}
\zeta^{(n)} &= \int w|p^{(n)}(a', dw) - wp^{(n)}(a, dw)| \\
&= \int wp^{(n-1)}(y, dw)|p(a', dy) - p(a, dy)| \\
&\le \int wp^{(n-1)}(y, dw)|p(a', dy) - p(a, dy)| \\
&= \int |p(a', dy) - p(a, dy)| \int wp^{(n-1)}(y, dw) \\
&\le \int w_{max}|p(a', dy) - p(a, dy)| \\
&\le c_0 w_{max}|a' - a|
\end{aligned}
$$

by (6). Thus the first part of the lemma is true with $k_1 = c_0 w_{max}$.

ii) Follows trivially by use of Theorem 7 with $k_2 = 2c$. ∎

Lemma 7: Let $a' > a$. Then,

$$\sum_{n=1}^{\infty} n^k \zeta^{(n)} \le f_k(a' - a),$$

where $f_k(x) \to 0$ as $x \to 0$, for all k.

Proof: Without loss of generality, let $k_2 > 2k_1 w_{max}/\gamma$. Let $n_0 \ge 0$ be such that

$$k_2\gamma^{n_0} \ge k_1|a' - a| \ge k_2\gamma^{n_0+1}. \tag{23}$$

It is easy to check that

$$n_0 = \left\lfloor \frac{\log(\frac{k_1}{k_2}|a' - a|)}{\log \gamma} \right\rfloor. \tag{24}$$

Then

$$
\begin{aligned}
\sum_{n=1}^{\infty} n^k \zeta^{(n)} &= \sum_{n=1}^{n_0} n^k \zeta^{(n)} + \sum_{n=n_0+1}^{\infty} n^k \zeta^{(n)} \\
&\le \sum_{n=1}^{n_0} n^k k_1|a' - a| + \sum_{n=n_0+1}^{\infty} n^k k_2\gamma^n \\
&= \underbrace{k_1 P_k(n_0)|a' - a|}_{T_1} + \underbrace{k_2 F_k(n_0)}_{T_2}
\end{aligned}
$$

for appropriately defined P_k and F_k. Now, $\lim_{x \to 0} x(\log x)^n = \lim_{y \to \infty} (-y)^n e^{-y} = 0$ (for $y = -\log x$), for all n. Using the value of n_0 from (24) in the above equation, it follows that $T_1 \to 0$ as $|a' - a| \to 0$. Also, note that as $|a' - a| \to 0$, $n_0 \to \infty$, and hence, $T_2 \to 0$. Also, T_1 and T_2 are functions of $|a' - a|$ because n_0 is. The lemma follows. ∎

Now, we describe in brief the Athreya-Ney-Nummelin construction (see [22]) of pseudo atom to construct another Markov chain $\check{M}$ related to the Markov chain $M \approx < W_n >$. The construction is as follows: Let $\check{W} = W \times \{0, 1\}$, where, $W^0 = W \times \{0\}$ and $W^1 = W \times \{1\}$ are thought of as copies of W, equipped with copies $\mathcal{B}(W^0)$ and $\mathcal{B}(W^1)$ of the Borel sigma field $\mathcal{B}(W)$ of W. Let $\mathcal{B}(\check{W})$ denote the smallest sigma field containing sets of the form $A^0 := A \times \{0\}$, $A^1 := A \times \{1\}$; $A \in \mathcal{B}(W)$. Let x^0, w^0, a^0 denote elements in W^0 and x^1, w^1, a^1 denote elements in W^1. For each measure λ on W, we define a measure λ^* on $\check{W}$ as follows:

$$\lambda^*(A^0) = (1 - \delta)\lambda(A)$$
$$\lambda^*(A^1) = \delta\lambda(A)$$

We now define the Markov chain $\check{M}$ on $\check{W}$ by listing down the single step transition kernel $\check{p}$.

$$\check{p}(x^0, \cdot) = \frac{p^*(x, \cdot) - \delta\nu^*(\cdot)}{1 - \delta}$$
$$\check{p}(x^1, \cdot) = \nu^*(\cdot) \tag{25}$$

That is,

$$\check{p}(x^0, A^0) = p(x, A) - \delta\nu(A)$$
$$\check{p}(x^0, A^1) = \frac{\delta}{1 - \delta}[p(x, A) - \delta\nu(A)]$$
$$\check{p}(x^1, A^0) = (1 - \delta)\nu(A)$$
$$\check{p}(x^1, A^1) = \delta\nu(A)$$

Note that these probabilities are well defined since the Minorization Condition, (7) holds. It is easy to prove (see p. 104 of [22] for the second claim, the first follows easily by induction.) that

$$\check{p}^{(n)}(x^i, A^1) = \frac{\delta}{1 - \delta}\check{p}^{(n)}(x^i, A^0), \ i = 0, 1,$$
$$p^{(n)}(x, A) = (1 - \delta)\check{p}^{(n)}(x^0, A^0 \cup A^1) + \delta\check{p}^{(n)}(x^1, A^0 \cup A^1). \tag{26}$$

The Markov chain M can be recovered from $\check{M}$ as follows: Let the initial distribution on M be $\lambda(dw)$.

Let the initial distribution on $\check{M}$ be $\lambda^*(dw)$. Then it is easy to prove by use of (26) that

$$\int_W \lambda(dw) \Pr(w_n \in A_n, \dots, w_1 \in A_1 | w_0 = w)$$

$$= \int_{\check{W}} \lambda^*(d\bar{w}) \Pr(\bar{w}_n \in A_n^0 \cup A_n^1, \dots, \bar{w}_1 \in A_1^0 \cup A_1^1 | \bar{w}_0 = \bar{w}) \quad (27)$$

Thus if n is the smallest time at which the chain is in W^1, the behavior of the chain from time $(n + 1)$ onwards is independent of where the chain started. This is because, as can be checked from (25), $p(x^1, \cdot)$ is independent of x^1. This property about $\check{M}$ helps derive many useful properties about $\check{M}$, and corresponding properties about M can be proved by use of (26). The set W^1 is called the *Pseudo Atom*. Next we state a few definitions and observations.

We define a partial order, $\geq$, on $\check{W}$ as follows:

Definition 5—Partial Order, $\geq$, on $\check{W}$: $w_1^i \geq w_2^j$ if $i = j$ and $w_1 \geq w_2$. Note that if $i \neq j$, w_1^i and w_2^j are not related by the partial order $\geq$.

Definition 6—Stochastic Dominance, $\succ$, on $\check{W}$: Let $\check{p}$ and $\check{q}$ be probability measures on $\check{W}$. We say that $\check{p} \succ \check{q}$ if $\check{p}|_{W^0} \succ \check{q}|_{W^0}$ and $\check{p}|_{W^1} \succ \check{q}|_{W^1}$ for all $w^i \in \check{W}$.

Lemma 8: Let $\check{X}_1$ and $\check{X}_2$ be random variables in $\check{W}$. Let $\check{X}_1 \sim \check{p}$, $\check{X}_2 \sim \check{q}$. Let $\check{p} \succ \check{q}$. Then, there exist random variables $\check{X}_1'$ and $\check{X}_2'$ such that $\check{X}_1' \sim \check{p}$, $\check{X}_2' \sim \check{q}$, and $\check{X}_1' \geq \check{X}_2'$ a. s.

This is a trivial extension of Lemma 2.

Lemma 9: $\check{p}(x'^0, \cdot) \succ \check{p}(x^0, \cdot)$, $\check{p}(x'^1, \cdot) = \check{p}(x^1, \cdot)$, where $x' \geq x$.

This is easy to prove.

Recall that our Markov chain M is the packet arrival process. Since M can be recovered from $\check{M}$, we identify our packet arrival process with the chain $\check{M}$. By using the structure on $\check{M}$, we will relate different sample paths of the packet arrival process and derive useful results.

Let $< W_k >$ and $< W_k' >$ be packet arrival processes with $W_0 \sim (1 - \delta)\kappa(a^0) + \delta\kappa(a^1)$, $W_0' \sim (1 - \delta)\kappa(a'^0) + \delta\kappa(a'^1)$, where $\kappa(x)$ stands for the Dirac measure at x. For each sample path $< w_k^{i_k} >$ of the process $< W_k >$, we associate a corresponding sample path $< w'^{i_k'}_k >$ of the process $< W_k' >$ as follows:

- Case I: k is such that $i_r = 0, 0 \leq r \leq (k-1)$. Assume, inductively, that $w'^{i_r'}_r \geq w_r^{i_r}, 0 \leq r \leq (k-1)$, and $i_r' = i_r, 0 \leq r \leq (k - 1)$. Now, $\check{p}(w'^{i_{k-1}'}_{k-1}, \cdot) = \check{p}(w'^0_{k-1}, \cdot) \succ \check{p}(w^0_{k-1}, \cdot) = \check{p}(w_{k-1}^{i_{k-1}}, \cdot)$. By Lemma 8, we can have an association wherein $w'^{i_k'}_k > w_k^{i_k}$, which also implies that $i_k' = i_k$.
- Case II: There exists $k' < k$ such that $i_{k'} = 1$, i.e., $< W_m >$ has hit the pseudo-atom before k. In this case, associate, $w'^{i_k'}_k = w_k^{i_k}$.

Thus, to each path $< w_k{}^{i_k} >$, we associate a corresponding sample path $< w'_k{}^{i'_k} >$ such that $w'_k{}^{i'_k} \geq w_k{}^{i_k}$, and equality holds for $n \geq (n_0 + 1)$, where n_0 is the smallest integer for which $i_{n_0} = 1$.

Definition 7—Coupling Time, τ_c : The time when $< W_k >$ hits W^1 for the first time is called the coupling time, and is denoted by τ_c.

Note that by our construction, $< W'_k >$ also hits W^1 at this time and this is the first time the two processes 'meet' at W^1. The reason for this nomenclature is that $< W_k >$ and $< W'_k >$ get "coupled" at this time in the sense that for all future time, $W'_k = W_k$. We will need the following two lemmas:

Lemma 10:

$$\Pr(\tau_c = n) = \delta(1 - \delta)^n, \ n \geq 0.$$

Proof: Note that $W_0 \sim (1 - \delta)\kappa(a^0) + \delta\kappa(a^1)$. Thus,

$$\Pr(W_0 \in W^0) = 1 - \delta,$$
$$\Pr(W_0 \in W^1) = \delta.$$

From (25), it is easy to see that

$$\Pr(W_{k+1} \in W^0 \mid W_k \in W^0) = 1 - \delta,$$
$$\Pr(W_{k+1} \in W^1 \mid W_k \in W^0) = \delta.$$

Now,

$$\Pr(\tau_c = 0) = \Pr(W_0 \in W^1) = \delta.$$

Hence

$$
\begin{aligned}
&\Pr(\tau_c = n) \\
=\ &\Pr(W_0 \in W^0, W_1 \in W^0, \ldots, W_{n-1} \in W^0, W_n \in W^1) \\
=\ &\Pr(W_0 \in W^0)\Big[\prod_{k=0}^{n-2} \Pr(W_{k+1} \in W^0 \mid W_k \in W^0)\Big] \Pr(W_n \in W^1 \mid W_{n-1} \in W^0) \\
=\ &(1 - \delta)(1 - \delta)^{n-1}\delta \\
=\ &\delta(1 - \delta)^n.
\end{aligned}
$$

∎

Definition 8: $\check{p}_\tau^{(n)}(a^i, A^j) \triangleq \Pr(W_n \in A^j \mid W_0 = a^i, \tau_c = \tau)$.

Lemma 11: Given $\tau_c = \tau \geq 1$,

$$\check{p}_\tau^{(n)}(a^0, A^0) = \frac{p^{(n)}(a, A) - \delta\eta_n(A)}{(1 - \delta)^n}, \qquad 0 \leq n \leq \tau - 1,$$

$$\check{p}_\tau^{(n)}(a^0, A^1) = 0, \qquad 0 \leq n \leq \tau - 1,$$

$$\check{p}_\tau^{(n)}(a^0, A^0) = 0, \qquad n = \tau,$$

$$\check{p}_\tau^{(n)}(a^0, A^1) = \frac{p^{(n)}(a, A) - \delta\eta_n(A)}{(1 - \delta)^n}, \qquad n = \tau.$$

where η_n are positive measures satisfying

$$\int_W \frac{p^{(n)}(a, dw) - \delta\eta_n(dw)}{(1 - \delta)^n} = 1.$$

Proof: We first prove the lemma for $n = 1$.

Case 1: $\tau > 1$. This means that $W_1 \in W^0$. This implies that $\check{p}_\tau^{(1)}(a^0, A^1) = 0$.

$$
\begin{aligned}
\check{p}_\tau^{(1)}(a^0, A^0) &= \Pr(W_1 \in A^0 \mid W_0 = a^0, \tau_c = \tau) \\
&= \frac{\Pr(W_1 \in A^0, \tau_c = \tau \mid W_0 = a^0)}{\Pr(\tau_c = \tau \mid W_0 = a^0)} \\
&= \frac{\Pr(\tau_c = \tau \mid W_1 \in A^0, W_0 = a^0)\Pr(W_1 \in A^0 \mid W_0 = a^0)}{\Pr(\tau_c = \tau \mid W_0 = a^0)} \\
&= \frac{\Pr(\tau_c = \tau \mid W_1 \in A^0)\check{p}(a^0, A^0)}{\Pr(\tau_c = \tau \mid W_0 = a^0)} \\
&= \frac{(1 - \delta)^{\tau-1}[p(a, A) - \delta\nu(A)]}{(1 - \delta)^\tau} \\
&= \frac{p(a, A) - \delta\nu(A)}{1 - \delta}.
\end{aligned}
$$

Thus the claim is true in the case $\tau > 1$, $n = 1$, with $\eta_1 \equiv \nu$.

Case 2: $\tau = 1$. The proof is similar to Case 1, and indeed, $\eta_1 \equiv \nu$.

Let the lemma be true for $n = k$. We prove the lemma for $n = (k + 1)$.

Case 1: $\tau > (k+1)$. Clearly, $\check{p}_\tau^{(k+1)}(a^0, A^1) = 0$.

$$\check{p}_\tau^{(k+1)}(a^0, A^0)$$

$$= \int_{W^0} \check{p}_\tau^{(1)}(a^0, dy^0) \check{p}_\tau^{(k)}(y^0, A^0)$$

$$= \int_W \left[\frac{p(a, dy) - \delta\nu(dy)}{(1-\delta)} \cdot \frac{p^{(k)}(y, A) - \delta\eta_k(A)}{(1-\delta)^k} \right]$$

$$= \frac{p^{(k+1)}(x, A) - \delta\left[\eta_k(A) - \delta\eta_k(A) + \int_W \nu(dy)p^{(k)}(y, A)\right]}{(1-\delta)^{k+1}}.$$

Thus, with $\eta_{k+1}(A) = [\eta_k(A) - \delta\eta_k(A) + \int_W \nu(dy)p^{(k)}(y, A)]$, Case 1 follows.

Case 2: $\tau = (k+1)$. A similar proof can be given, and indeed, η_{k+1} is the same as in Case 1.
The lemma follows by induction. ∎

Next, we estimate $V^\beta(x, a', \mu) - V^\beta(x, a, \mu)$. Let $< Y_k >$ be optimal for the process $< X_k >$, defined inductively as:

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, W_0 \sim (1-\delta)\kappa(a^0) + \delta\kappa(a^1), \mu_0 = \mu.$$

Consider the process $< X'_k >$ defined inductively, for $0 \le k \le \tau_c$ as:

$$X'_{k+1} = X'_k - Y_k + W'_{k+1}, X_0 = x, W'_0 = W_0 \sim (1-\delta)\kappa(a^0) + \delta\kappa(a^1), \ 0 \le k \le \tau_c, \mu_0 = \mu$$

where, to each sample path $< w_k^{i_k} >$ of $< W_k >$ we have associated the sample path $< w'^{i'_k}_k >$ of $< W'_k >$ as before. The channel fading process $< \mu_n >$ is the same for both. Since $w'_k > w_k$, the policy $< Y_k >$ is admissible in the process $< X'_k >$. Note also, that $< Y_k >$ need not be optimal for $< X'_k >$. Hence, for $k \ge 0$,

$$E\left[V^\beta(X'_k, W'_k, \mu_k) - V^\beta(X_k, W_k, \mu)\right] \le E\left[\lambda(X'_k - X_k) \right.$$
$$\left. + \beta\left\{V^\beta(X'_{k+1}, W'_{k+1}, \mu_{k+1}) - V^\beta(X_{k+1}, W_{k+1}, \mu_{k+1})\right\}\right].$$

Iterating over $0 \le k \le \tau_c$ and taking expectations,

$$V^\beta(x, a', \mu) - V^\beta(x, a, \mu) \le E\left[\lambda\sum_{k=0}^{\tau_c} \beta^k(X'_k - X_k) \right.$$
$$\left. + \beta^{\tau_c+1}\left\{V^\beta(X'_{\tau_c+1}, W'_{\tau_c+1}, \mu_{\tau_c+1}) - V^\beta(X_{\tau_c+1}, W_{\tau_c+1}, \mu_{\tau_c+1})\right\}\right]. \quad (28)$$

But

$$
E\left[\sum_{k=0}^{\tau_c} \beta^k (X_k' - X_k)\right]
$$

$$
\leq \quad E\left[\sum_{k=0}^{\tau_c} (X_k' - X_k)\right]
$$

$$
= \quad E\left[\sum_{k=1}^{\tau_c} (\tau_c - k + 1)(W_k' - W_k)\right]
$$

$$
= \quad \sum_{n=1}^{\infty} p(\tau_c = n) E\left[\sum_{k=1}^{n} (n - k + 1)(W_k' - W_k) \ \Big| \ \tau_c = n\right]
$$

$$
=^{*1} \quad \sum_{n=1}^{\infty} \delta(1-\delta)^n \sum_{k=1}^{n} (n - k + 1) \int_{W^0} w \left[\check{p}_{\tau_c}^{(k)}(a'^0, dw^0) - \check{p}_{\tau_c}^{(k)}(a^0, dw^0)\right]
$$

$$
=^{*2} \quad \sum_{n=1}^{\infty} \delta(1-\delta)^n \sum_{k=1}^{n} (n - k + 1)\frac{\zeta^{(k)}}{(1-\delta)^k}
$$

$$
= \quad \sum_{k=1}^{\infty} \zeta^{(k)} \sum_{n=k}^{\infty} (n - k + 1)\delta(1-\delta)^{n-k}
$$

$$
= \quad \sum_{k=1}^{\infty} \zeta^{(k)} \left(\sum_{n=0}^{\infty} (n+1)\delta(1-\delta)^n\right)
$$

$$
=^{*3} \quad O\left(\sum_{n=1}^{\infty} \zeta^{(n)}\right)
$$

$$
=^{*4} \quad O[f_0(|a' - a|)] \to 0 \text{ as } a' \to a. \tag{29}
$$

Here

- $(*^1)$ follows from Lemma 10.
- $(*^2)$ follows from Definition 4 and Lemma 11.
- $(*^3)$ follows by interchange of summations.
- $(*^4)$ follows by Lemma 7.

Thus for some $K > 0$,

$$E[V^\beta(X'_{\tau_c+1}, W'_{\tau_c+1}, \mu_{\tau_c+1}) - V^\beta(X_{\tau_c+1}, W_{\tau_c+1}, \mu_{\tau_c+1})]$$

$$=^{*1} \quad E[V^\beta(X'_{\tau_c+1}, W_{\tau_c+1}, \mu_{\tau_c+1}) - V^\beta(X_{\tau_c+1}, W_{\tau_c+1}, \mu_{\tau_c+1})]$$

$$= \quad \sum_{n=0}^{\infty} p(\tau_c = n) E\left[V^\beta\left(x - \sum_{k=0}^{n} Y_k + \sum_{k=1}^{n} W'_k + W_{n+1},\ W_{n+1}, \mu_{n+1}\right)\right.$$

$$\left. - V^\beta\left(x - \sum_{k=0}^{n} Y_k + \sum_{k=1}^{n+1} W_k,\ W_{n+1}, \mu_{n+1}\right) \middle| \tau_c = n\right]$$

$$\leq^{*2} \quad K \sum_{n=1}^{\infty} p(\tau_c = n) E\left[\left(\sum_{k=1}^{n} W'_k - \sum_{k=1}^{n} W_k\right)\right.$$

$$\left. + \left(\sum_{k=1}^{n} W'_k + \sum_{k=1}^{n} W_k\right)\left(\sum_{k=1}^{n} W'_k - \sum_{k=1}^{n} W_k\right) \middle| \tau_c = n\right]$$

$$\leq^{*3} \quad K \sum_{n=1}^{\infty} p(\tau_c = n) \left[\sum_{k=1}^{n} \frac{\zeta^{(k)}}{(1-\delta)^k} + 2n w_{max} \sum_{k=1}^{n} \frac{\zeta^{(k)}}{(1-\delta)^k}\right]$$

$$=^{*4} \quad O\left(\sum_{n=0}^{\infty} n \zeta^{(n)}\right)$$

$$=^{*5} \quad O[f_0(|a'-a|) + f_1(|a'-a|)] \to 0 \text{ as } a' \to a. \tag{30}$$

- $(*^1)$ follows because for $n > \tau_c$, $W'_n = W_n$.
- $(*^2)$ follows by Lemma 5.
- $(*^3)$ Here, we have used the fact that if X and Y are non-negative random variables and the maximum possible value of X is x_{max}, $E(XY) \leq x_{max} EY$. In this case, $X = [\sum_{k=1}^{n} W'_k + \sum_{k=1}^{n} W_k]$ and $Y = [\sum_{k=1}^{n} W'_k - \sum_{k=1}^{n} W_k]$.
- $(*^4)$ follows by interchanging summations, as in (29) above.
- $(*^5)$ follows by Lemma 7.

Using (29) and (30) in (28), we get:

$$V^\beta(x, a, \mu') - V^\beta(x, a, \mu) \leq O(f_0(|a'-a|) + f_1(|a'-a|)). \tag{31}$$

It follows by Theorem 5, Lemma 5 and (31) that

$$V^\beta(x', a', \mu') - V^\beta(x, a, \mu)$$

$$\leq \quad k_1[|x'-x| + |x'-x|(x'-x)] + k_2(f_0(|a'-a|) + f_1(|a'-a|))$$

$$+ |\mu' - \mu| F(\min(y_{max}, x)). \tag{32}$$

Definition 9—$\bar{\mathbf{V}}^\beta$: $\bar{V}^\beta(x, a, \mu) = V^\beta(x, a, \mu) - V^\beta(0, 0, 0)$.

Lemma 12:   i) For all x, a, μ, there exists M depending on x, a, μ such that $\bar{V}^\beta(x, a, \mu) \leq M$ for all β.

ii) $\bar{V}^\beta, \beta \in (0, 1)$, is an equicontinuous family of functions.

Both parts of the lemma follow from (32).

Lemma 13: There exists M such that $(1 - \beta)V^\beta(0, 0, 0) \leq M$ for all β.

   Proof:

$$
\begin{aligned}
V^\beta(0, 0, 0) \ =^{*^1} \ & \beta \int q(d\nu)p(0, dw)V^\beta(w, w, \nu) \\
\leq \ & \beta V^\beta(w_{max}, w_{max}, \mu_{max}) \\
\leq^{*^2} \ & \beta[V^\beta(0, 0, 0) + k_1(w_{max} + w_{max}^2) + k_2(f_0(w_{max}) + f_1(w_{max})) \\
& \qquad\qquad\qquad\qquad\qquad\qquad + \mu_{max}F(\min(w_{max}, y_{max}))]
\end{aligned}
$$

$$
\Rightarrow (1-\beta)V^\beta(0,0,0) \leq \beta[k_1(w_{max}+w_{max}^2)+k_2(f_0(w_{max})+f_1(w_{max}))+\mu_{max}F(\min(w_{max}, y_{max}))]
$$

Here, k_1 and k_2 are some constants. $(*^1)$ follows by the Bellman equation and $(*^2)$ follows by (32). The lemma follows.   ■

   By Lemma 12, $\bar{V}^\beta$ is a bounded equicontinuous family of functions. By the Arzela-Ascoli Theorem (see p. 214 of [19]), there exists a sequence $< \beta_n' > \uparrow 1$ such that $\lim_{n\to\infty} \bar{V}^{\beta_n'} = \bar{V}$, and the convergence is uniform on compacts. By Lemma 13, $(1-\beta)V^\beta(0, 0, 0)$ is bounded. Hence, there exists a subsequence $< \beta_n >$ of $< \beta_n' >$, $\beta_n \uparrow 1$, such that $\lim_{n\to\infty}(1 - \beta_n)V^{\beta_n}(0, 0, 0) = \rho$, for some ρ. Consider the Bellman equation, (15),

$$
V^\beta(x, a, \mu) = \min_{y \leq x \wedge y_{max}} [\lambda x + \mu F(y) + \beta \int q(d\nu)p(a, dw)V^\beta(x - y + w, w, \nu)].
$$

After some manipulation and using Definition 9,

$$
\begin{aligned}
\bar{V}^\beta(x, a, \mu) = \min_{y \leq x \wedge y_{max}} [\lambda x + \mu F(y) - (1 - \beta)\bar{V}^\beta(0, 0, 0) \\
+ \beta \int q(d\nu)p(a, dw)\bar{V}^\beta(x - y + w, w, \nu)].
\end{aligned}
$$

Letting $\beta \to 1$ on both sides along β_n, we get:

$$
\bar{V}(x, a, \mu) = \min_{y \leq x \wedge y_{max}} [\lambda x + \mu F(y) - \rho + \int q(d\nu)p(a, dw)\bar{V}(x - y + w, w, \nu)]. \tag{33}
$$

This follows by arguments similar to those used to prove the convergence of V_n^β to V^β in Section V, except that in place of the Monotone Convergence Theorem, we have to use the Dominated Convergence Theorem (see p. 381 of [19]). The applicability of the Dominated Convergence Theorem follows by use of (32). The uniform convergence on compact sets follows by equicontinuity. (33) is the Bellman equation for the Average Cost problem. Let $\mathbf{v} = (x, a, \mu)$ denote the state. Define:

$$y^*(\mathbf{v}) = \arg \min_{y \leq x \wedge y_{max}} [\lambda x + \mu F(y) - \rho + \int q(d\nu)p(a, dw)\bar{V}(x - y + w, w, \nu)]. \tag{34}$$

Note that y^* is a stationary policy. Recall that we want to minimize $C_\lambda = P + \lambda D$, where P and D are defined in Definition 1.

Theorem 8: The ρ in (33) equals the optimal C_λ, attained when policy y^* is used.

Proof: Let ρ_{opt} be the infimum of all achievable average costs. Since the chain is ergodic under y^*, ρ_{opt} is independent of the initial state. Also, it is known that stable stationary policies suffice for optimality [23]. We will first prove that $\bar{V}$ is bounded from below. We have

$$(1 - \beta) \inf V^\beta \leq (1 - \beta)E[\sum_{n=0}^{\infty} \beta^n (\mu_n F(y_n) + \lambda x_n)],$$

where the expectation on the right is under some stable stationary policy π and an arbitrary deterministic initial state. Averaging both sides w.r.t. the stationary distribution under π and then taking infimum over all such π, we obtain, in view of the above observation from [23], that $(1 - \beta) \inf V^\beta \leq \rho_{opt}$. It is easy to see that $V^\beta(x, a, \mu) \uparrow \infty$ as $x \uparrow \infty$ and hence this infimum is a minimum attained at some $(x_\beta, a_\beta, \mu_\beta)$. Let $y^\beta(\cdot)$ denote the optimal stationary policy for the β−discounted problem. In view of the dynamic programming equation (15) above, we have

$$0 = \lambda x + \mu F(y^\beta(x, a, \mu)) + \beta \int q(d\nu)p(a, dw)V^\beta(x - y + w, w, \nu) - V^\beta(x, w, \mu).$$

Therefore, at $(x_\beta, a_\beta, \mu_\beta)$, we have:

$$0 \geq \lambda x_\beta + \beta V^\beta(x_\beta, a_\beta, \mu_\beta) - V^\beta(x_\beta, a_\beta, \mu_\beta),$$

that is,

$$\lambda x_\beta \leq (1 - \beta)V^\beta(x_\beta) \leq \rho_{opt}.$$

That is, V^β and hence $\bar{V}^\beta$ attain their minimum on the bounded set $\{||x|| \leq \frac{\rho_{opt}}{\lambda}\}$. Then so will $\bar{V}$. Now, let $\mathbf{v}_n = (x_n, w_n, \mu_n)$ denote the state at the beginning of the nth slot. Also, let $c_n(\mathbf{v}_n, y_n) = [\lambda x_n + \mu_n F(y_n)]$ be the nth step cost. From the Bellman equation, it follows that

$$\bar{V}(\mathbf{v}_n) = \min_{y_n \leq x_n \wedge y_{max}} [c_n(\mathbf{v}_n, y_n) - \rho + E[\bar{V}(\mathbf{v}_{n+1})|\mathbf{v}_n]]. \tag{35}$$

Thus

$$\bar{V}(\mathbf{v}_n) = c(\mathbf{v}_n, y^*(\mathbf{v}_n)) - \rho + E[\bar{V}(\mathbf{v}_{n+1})|\mathbf{v}_n].$$

Taking expectation of both sides given the initial state $\mathbf{v}_0$,

$$E[\bar{V}(\mathbf{v}_n)|\mathbf{v}_0] = E[c(\mathbf{v}_n, y^*(\mathbf{v}_n))|\mathbf{v}_0] - \rho + E[\bar{V}(\mathbf{v}_{n+1})|\mathbf{v}_0]$$

$$\Rightarrow \quad \sum_{n=0}^{N-1} E[\bar{V}(\mathbf{v}_n)|\mathbf{v}_0] = \sum_{n=0}^{N-1} E[c(\mathbf{v}_n, y^*(\mathbf{v}_n))|\mathbf{v}_0] - N\rho + \sum_{n=0}^{N-1} E[\bar{V}(\mathbf{v}_{n+1})|\mathbf{v}_0]$$

$$\Rightarrow \quad \frac{1}{N} \sum_{n=0}^{N-1} E[c(\mathbf{v}_n, y^*(\mathbf{v}_n))|\mathbf{v}_0] = \rho + \frac{\bar{V}(\mathbf{v}_0) - E[\bar{V}(\mathbf{v}_N)|\mathbf{v}_0]}{N}$$

$$\Rightarrow^{*^1} \quad \limsup_{N\to\infty} \frac{1}{N} E\Big[\sum_{n=0}^{N-1} c(\mathbf{v}_n, y^*(\mathbf{v}_n))|\mathbf{v}_0\Big] \leq \rho$$

$$\Rightarrow \quad \rho_{opt} \leq \rho.$$

Here $(*^1)$ uses Fatou's lemma (Chapter 6 of [19]) and the fact that $\bar{V}(\cdot)$ is bounded from below. Next we prove that $\rho \leq \rho_{opt}$. Since stationary policies suffice for optimality, there exists a stationary policy which attains the optimal average cost. (This can be proved *a priori* by the general techniques of [23]. Even otherwise, one could argue with an '$\epsilon-$optimal stationary policy' instead with slight additional work.) Let π_0 be the stationary policy which attains the optimal average cost for the average cost problem. Now for a stationary policy π,

$$\liminf_{N\to\infty} \frac{1}{N} E \sum_{n=0}^{N-1} c(\mathbf{v}_n, y^\pi(\mathbf{v}_n)) = \limsup_{N\to\infty} \frac{1}{N} E \sum_{n=0}^{N-1} c(\mathbf{v}_n, y^\pi(\mathbf{v}_n)) \tag{36}$$

By the Tauberian theorem (see [30]), and (36), it follows that for any stationary policy π,

$$\lim_{N\to\infty} \frac{1}{N} E \sum_{n=0}^{N-1} c(\mathbf{v}_n, y^\pi(\mathbf{v}_n)) = \lim_{\beta\uparrow 1}(1-\beta) E \sum_{n=0}^{\infty} \beta^n c(\mathbf{v}_n, y^\pi(\mathbf{v}_n)) \tag{37}$$

Now,

$$V^\beta(0,0,0) \leq J_{\pi_0}^\beta(0,0,0)$$

$$\Rightarrow \quad \lim_{\beta\uparrow 1}(1-\beta)V^\beta(0,0,0) \leq \lim_{\beta\uparrow 1}(1-\beta)J_{\pi_0}^\beta(0,0,0)$$

$$\Rightarrow^* \quad \rho \leq \rho_{opt}$$

(*) follows from the definitions of ρ, ρ_{opt} and π_0 and from (37). This proves the theorem. ∎

Theorem 9: i) (33) uniquely specifies ρ as ρ_{opt}. Also, if (V, ρ_{opt}) is another continuous solution to (33) satisfying $E_s[|V(X_n) - \bar{V}(X_n)|] < \infty$, where $E_s[\,\cdot\,]$ is the stationary average under any

optimal strategy, then V differs from $\bar{V}$ by at most an additive scalar on the support the stationary law of any optimal stationary policy.

ii) $\bar{V}(x, a, \mu)$ is continuous increasing in a, μ and continuous convex increasing in x.

iii) For all $\mu, x' \geq x, a' \geq a$,

$$\bar{V}(x', a', \mu) + \bar{V}(x, a, \mu) \geq \bar{V}(x, a', \mu) + \bar{V}(x', a, \mu) \tag{38}$$

(This is the 'supermodularity' property for $\bar{V}$.)

Proof: Uniqueness of ρ is proved as above. Let (V, ρ_{opt}) be another solution to (33) with V continuous and $E_s[|V(X_n) - \bar{V}(X_n)|] < \infty$. Consider the process governed by the stationary policy y^* above. Since this policy in particular has finite cost, it follows that it is stable. Consider the stationary $< X_n >$ governed by y^*. By the definition of y^*,

$$0 = -\bar{V}(X_n) + \lambda X_n + \mu_n F(y^*(X_n, W_n, \mu_n)) - \rho_{opt} + E[\bar{V}(X_{n+1})|X_n].$$

Also, since (33) is satisfied by (V, ρ_{opt}),

$$0 \leq -V(X_n) + \lambda X_n + \mu_n F(y^*(X_n, W_n, \mu_n)) - \rho_{opt} + E[V(X_{n+1})|X_n].$$

From above, it follows that $\bar{V}(X_n) - V(X_n), n \geq 0$, is a supermartingale with $\sup_m E_s[|V(X_m) - \bar{V}(X_m)|] = E_s[|V(X_n) - \bar{V}(X_n)|] < \infty$. Thus it must converge a.s. by the supermartingale convergence theorem. Since $< X_n >$ is ergodic, this is possible only if $V - \bar{V}$ is constant a.s. w.r.t. the stationary law of X_n. By continuity, it is so on the support thereof. The first claim follows. Using Definition 9 it is easy to see that the second and the third claims hold for $\bar{V}^\beta$. Let $\beta \to 1$ along β_n to conclude. ∎

In order to prove the structural results, we need the following technical result:

Theorem 10: Let h be a function of $(n + 2)$ variables, viz., scalars y, x, and n-dimensional vector $\mathbf{s}$. Let S be a sublattice of $\mathbb{R}^n$. The sublattice structure defines a partial order $>$ on S. Let h be defined on the set (x, y, S), where $x \in \mathbb{R}$, and $0 \leq y \leq x \wedge y_{max}$. Let h be a continuous function. Let h satisfy decreasing differences in $(y, x, \mathbf{s})$, that is, for $x' \geq x, \mathbf{s}' \geq \mathbf{s}, y' \geq y$, and $y, y' \leq x$:

$$h(y', x', \mathbf{s}') - h(y, x', \mathbf{s}') \leq h(y', x, \mathbf{s}) - h(y, x, \mathbf{s}). \tag{39}$$

Let

$$D^*(x, \mathbf{s}) = \arg\min\{h(y, x, \mathbf{s}) | y \in [0, x]\}.$$

Then,

- For each $(x, \mathbf{s})$, $D^*(x, \mathbf{s})$ is a non-empty compact sublattice of R, and admits a greatest element denoted by $y^*(x, \mathbf{s})$.

- $y^*(x', \mathbf{s}') \geq y^*(x, \mathbf{s})$ whenever $x' \geq x$, $\mathbf{s}' \geq \mathbf{s}$.

Proof: This follows along the lines of Theorems 10.7 and 10.12 of [31]. ∎

We now prove structural results about the optimal stationary policies. Note that $y(x, a, \mu)$ denotes the number of packet transmissions in state (x, a, μ). Note that since the policy is stationary, the action taken depends only on the present state.

Theorem 11: There exists an optimal stationary policy such that $y^*(x', a', \mu') \geq y^*(x, a, \mu)$, for all $x' \geq x$, $a' \geq a$, $\mu' \leq \mu$.

Proof: Let $\chi = -\mu$, $\chi' = -\mu'$. Thus, $x' \geq x$, $a' \geq a$, $\chi' \geq \chi$.

Let $y' \geq y$, and $y, y' \leq x$. Lemma 14 (Appendix - C) and (38) imply that

$$\left[\bar{V}(x' - y + w, w, \mu) - \bar{V}(x' - y' + w, w, \mu) \right] \uparrow \text{ as } w \uparrow \forall \mu. \tag{40}$$

Lemma 1 and (40) imply

$$\int q(d\mu)p(a', dw) \left[\bar{V}(x' - y + w, w, \mu) - \bar{V}(x' - y' + w, w, \mu) \right]$$
$$\geq \int q(d\mu)p(a, dw) \left[\bar{V}(x' - y + w, w, \mu) - \bar{V}(x' - y' + w, w, \mu) \right]. \tag{41}$$

By convexity,

$$\bar{V}(x' - y + w, w, \mu) - \bar{V}(x' - y' + w, w, \mu)$$
$$\geq \bar{V}(x - y + w, w, \mu) - \bar{V}(x - y' + w, w, \mu), \quad \forall \mu. \tag{42}$$

Using (42) in (41) we get

$$\int q(d\mu)p(a', dw) \left[\bar{V}(x' - y + w, w, \mu) - \bar{V}(x' - y' + w, w, \mu) \right]$$
$$\geq \int q(d\mu)p(a, dw) \left[\bar{V}(x - y + w, w, \mu) - \bar{V}(x - y' + w, w, \mu) \right]. \tag{43}$$

Define

$$h(y, x, a, \mu) = \lambda x + \mu F(y) + \int q(d\nu)p(a, dw)\bar{V}(x - y + w, w, \nu).$$

Set $\bar{h}(\,\cdot\,,\,\cdot\,,\,\cdot\,,\,\cdot\,) = h(\,\cdot\,,\,\cdot\,,\,\cdot\,,\,-\,\cdot\,)$. Manipulating (43), we get

$$\bar{h}(y', x', a', \chi') - \bar{h}(y, x', a', \chi') \leq \bar{h}(y', x, a, \chi) - \bar{h}(y, x, a, \chi), \tag{44}$$

where in the notation of Theorem 10, $\mathbf{s} = (a, \chi)$, and $\bar{h}$ satisfies decreasing differences. By Theorem 10, the claim follows. ∎

This completes the proofs of Theorems 1 and 2. By Theorem 2 and the discussion that follows it, the same holds for the constrained problem when F is strictly convex.

VII. CONCLUSIONS

We have derived structural properties of the optimal transmission policy across a randomly varying channel for a single transmitter. By casting it as a constrained Markov decision process in discrete time with time-averaged costs, we prove structural results about the dependence of the optimal policy on buffer occupancy, number of packet arrivals in the previous slot and the channel fading for both i.i.d. and Markovian arrivals and channel states. When the packet arrival process is FSD Markovian and the channel fading is i.i.d., there exists an optimal stationary deterministic policy which is increasing in buffer occupancy and number of packet arrivals in the previous slot, and decreasing in the channel fading state. When the packet arrival process is FSD Markovian and the channel fading is Markovian, there exists an optimal stationary deterministic policy which is increasing in buffer occupancy and number of packet arrivals in the previous time slot. Nothing can, in general, be said about the nature of the optimal policy with respect to channel fading.

The main contribution is the methodology developed for the average cost criterion for continuous state spaces, which, while being the preferred one in communications applications, is notoriously difficult to handle rigorously. This methodology combines a pathwise comparison based on stochastic dominance with 'coupling at pseudo atom' to establish boundedness of the renormalized discounted value function, uniformly in the discounted factor as it tends to zero. It also uses a novel argument to show that the discounted value function itself attains its minimum on a bounded set independent of the discount factor, a fact that plays a key role in analyzing the dynamic programming equation for the ergodic problem. We expect the techniques developed here to have much broader implications.

VIII. ACKNOWLEDGEMENTS

APPENDIX

A. Proof of Theorem 3

Let $x < x'$. Let the arrival process $< W_k >$ be defined on some probability space $(\Omega, \mathcal{F}, P)$. Let $< Y'_k >$ be optimal for $< X'_k >$ given inductively by

$$X'_{k+1} = X'_k - Y'_k + W_{k+1}, \ X'_0 = x', W_0 = a, 0 \le k \le (n-1). \tag{45}$$

In particular, $Y'_k \le y_{max} \ \forall k$. Construct inductively, with the same channel fading process,

$$X_{k+1} = X_k - Y_k + W_{k+1}, \ X_0 = x, W_0 = a, 0 \le k \le (n-1), \tag{46}$$

where Y_k is the process $\min(X_k, Y'_k)$. Note that $Y_k \le Y'_k \le y_{max} \ \forall k$. It is easy to prove that $X_k \le X'_k$. The channel state process $\{\mu_n, n \ge 1\}$ is the same for both, with $\mu_0 = \mu$. "With the same channel fading process" means that we take the same realization of the random variable of channel fading for both X' and X.

Thus

$$V_n(x, a, \mu) \le J_{n,<Y_k>}(x, a, \mu) = E\left(\sum_{i=0}^{n-1} \beta^i[\lambda X_i + \mu_n F(Y_i)] + \lambda\beta^n X_n\right)$$
$$\le E\left(\sum_{i=0}^{n-1} \beta^i[\lambda X'_i + \mu_n F(Y'_i)] + \lambda\beta^n X'_n\right) = V_n(x', a, \mu) \tag{47}$$

Next we prove the continuity of V_n. Consider the same process $< W_k >$ with $W_0 = a$. Let $< Y_k >$ be optimal for $< X_k >$ given by

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, 0 \le k \le (n-1). \tag{48}$$

Consider the process

$$X'_{k+1} = X'_k - Y_k + W_{k+1}, X_0 = x', 0 \le k \le (n-1). \tag{49}$$

It is easy to check that $(X'_k - X_k) = (x' - x)$. Also, the transmission process $< Y_k >$ is the same in both cases. By using the fact that $V_n(x', a, \mu) \ge V_n(x, a, \mu)$, it is easy to see that

$$0 \le V_n(x', a, \mu) - V_n(x, a, \mu) \le J_{n,<Y_k>}(x', a, \mu) - V_n(x, a, \mu) = \sum_{i=0}^{n} \beta^i \lambda(x' - x)$$
$$= \frac{\lambda(1 - \beta^{n+1})}{1 - \beta}(x' - x) \le \frac{\lambda}{1 - \beta}(x' - x) \tag{50}$$

This proves the continuity of V_n and derives the required bound. Convexity is proved by induction. $V_0(x, a, \mu) = \lambda x$ is a convex function. Assume that $V_{n-1}(x, a, \mu)$ is convex $\forall a, \mu$. For $x = x_1, x_2$, let the minimum in (9) be attained, respectively, at $y = y_1, y_2$. Then

$$
\begin{aligned}
V_n(x_1, a, \mu) &= \lambda x_1 + \mu F(y_1) + \beta \int q(d\nu)p(a, dw)V_{n-1}(x_1 - y_1 + w, w, \nu) \\
V_n(x_2, a, \mu) &= \lambda x_2 + \mu F(y_2) + \beta \int q(d\nu)p(a, dw)V_{n-1}(x_2 - y_2 + w, w, \nu) \\
V_n(x_1, a, \mu) + V_n(x_2, a, \mu) &= \lambda(x_1 + x_2) + \mu(F(y_1) + F(y_2)) \\
&\quad + \beta \int q(d\nu)p(a, dw)[V_{n-1}(x_1 - y_1 + w, w, \nu) \\
&\quad + V_{n-1}(x_2 - y_2 + w, w, \nu)] \\
&\geq^{*1} 2\Big[\lambda\Big(\frac{x_1 + x_2}{2}\Big) + \mu F\Big(\frac{y_1 + y_2}{2}\Big) \\
&\quad + \beta \int q(d\nu)p(a, dw)V_{n-1}\Big(\frac{x_1 + x_2}{2} - \frac{y_1 + y_2}{2} + w, w, \nu\Big)\Big] \\
&= 2h_n\Big(\frac{y_1 + y_2}{2}, \frac{x_1 + x_2}{2}, a, \mu\Big) \\
&\geq^{*2} 2V_n\Big(\frac{x_1 + x_2}{2}, a, \mu\Big)
\end{aligned}
$$

$(*^1)$ follows by the convexity of V_{n-1} and F. $(*^2)$ follows since $h_n(y, x, a, \mu) \geq V_n(x, a, \mu) \ \forall x$. This proves the convexity of V_n assuming the convexity of V_{n-1}. By induction, V_n is convex for all n.

B. Proof of Theorem 4

Without loss of generality, let $a' \geq a$. We first prove that $V_n(x, a', \mu) \geq V_n(x, a, \mu)$. Let $< W_k >$ and $< W'_k >$ denote arrival processes corresponding to a and a' arrivals before slot zero. Thus, $W_0 = a$ and $W'_0 = a'$. To each sample path $< \tilde{w}_k >$ of $< W_k >$, we associate a sample path $< w'_k >$ of $< W'_k >$ as follows:

$w'_0 = a' \geq a = \tilde{w}_0$. Assume inductively that $w'_k \geq \tilde{w}_k$. Now, $w'_{k+1} \sim p(w'_k, dw) \succ p(\tilde{w}_k, dw) \sim \tilde{w}_{k+1}$. By Lemma 2, we can have an association wherein $w'_{k+1} \geq \tilde{w}_{k+1} \ a.s.$

Thus, we have arrival sequences $< W_k >$ and $< W'_k >$ such that $W_0 = a$, $W'_0 = a'$, and $W'_k \geq W_k$ $a.s.$ Let $< Y'_k >$ be optimal for $< X'_k >$, defined inductively by

$$
X'_{k+1} = X'_k - Y'_k + W'_{k+1}, X'_0 = x, W'_0 = a', 0 \leq k \leq (n-1). \tag{51}
$$

Construct inductively, with the same channel fading process,

$$
X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, W_0 = a, 0 \leq k \leq (n-1), \tag{52}
$$

where $<Y_k>= \min(X_k, Y_k') \le Y_k'$. Also, it is easy to see that $X_k \le X_k'$. It follows that

$$V_n(x, a, \mu) \le J_{n,<Y_k>}(x, a, \mu) = E\left(\sum_{i=0}^{n-1} \beta^i[\lambda X_i + \mu_i F(Y_i)] + \lambda \beta^n X_n\right)$$

$$\le E\left(\sum_{i=0}^{n-1} \beta^i[\lambda X_i' + \mu_i F(Y_i')] + \lambda \beta^n X_n'\right) = V_n(x, a', \mu). \quad (53)$$

We next derive the given bound. Consider the same processes $<W_k>, <W_k'>$. Let $<Y_k>$ be optimal for $<X_k>$ given inductively by

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, W_0 = a, 0 \le k \le (n-1) \quad (54)$$

Consider the process

$$X_{k+1}' = X_k' - Y_k + W_{k+1}', X_0' = x, W_0' = a', 0 \le k \le (n-1) \quad (55)$$

It is easy to check that

$$(X_k' - X_k) = \sum_{i=1}^{k}(W_i' - W_i) \quad (56)$$

Denoting by $p^{(k)}(a, dw)$ the k-step transition kernel for $<W_n>$, we have,

$$\begin{aligned} E(W_k' - W_k) &= \int [w p^{(k)}(a', dw) - w p^{(k)}(a, dw)] \\ &= \int\int w p^{(k-1)}(y, dw)[p(a', dy) - p(a, dy)] \\ &\le \int\int w p^{(k-1)}(y, dw)|p(a', dy) - p(a, dy)| \\ &= \int |p(a', dy) - p(a, dy)| \int w p^{(k-1)}(y, dw) \end{aligned}$$

$$(57)$$

$$\begin{aligned} &\le \int w_{max}|p(a', dy) - p(a, dy)| \\ &= w_{max} \int |p(a', dy) - p(a, dy)| \end{aligned} \quad (58)$$

Recall that w_{max} is the maximum number of packet arrivals in a slot. Note that the transmission process $<Y_k>$ is the same in both cases. It is easy to see by use of (56) and (57) that

$$\begin{aligned} V_n(x, a', \mu) - V_n(x, a, \mu) &\le J_{n,<Y_k>}(x, a', \mu) - V_n(x, a, \mu) \\ &\le \frac{w_{max}\lambda\beta}{(1-\beta)^2} \int |p(a', dy) - p(a, dy)| \end{aligned} \quad (59)$$

This completes the proof.

C. Proof of Theorem 6

To prove this Theorem, we first prove the following technical Lemma.

Lemma 14: Let $f(x, a)$ be a function which is convex in x for each a, and satisfies

$$f(x + \Delta x, a + \Delta a) + f(x, a) \geq f(x, a + \Delta a) + f(x + \Delta x, a), \tag{60}$$

where Δx and Δa are arbitrary non-negative real numbers. Then,

$$f(x + \Delta x + w, w) - f(x + w, w)$$

increases with w.

Proof: Let $\Delta w > 0$. Consider

$$\Big[f(x + \Delta x + w + \Delta w, w + \Delta w) - f(x + w + \Delta w, w + \Delta w)\Big]$$
$$- \Big[f(x + \Delta x + w, w) - f(x + w, w)\Big]$$
$$= \Big[f(x + \Delta x + w + \Delta w, w + \Delta w) - f(x + \Delta x + w, w + \Delta w)$$
$$- f(x + w + \Delta w, w + \Delta w) + f(x + w, w + \Delta w)\Big]$$
$$+ \Big[f(x + \Delta x + w, w + \Delta w) - f(x + w, w + \Delta w)$$
$$- f(x + \Delta x + w, w) + f(x + w, w)\Big]$$
$$\geq^* 0 + 0$$
$$= 0,$$

where (*) follows by the convexity of $f(\cdot, w)$ and (60). ∎

We can now prove Theorem 6 as follows.

Proof: $V_0(x, a, \mu) = \lambda x$ satisfies the assertion. Assume that the assertion is true for V_{n-1}. We will prove that the assertion is true for V_n. Consider the Bellman equation,

$$V_n(x, a, \mu) = \min_{y \leq x \wedge y_{max}} [\lambda x + \mu F(y) + \beta \int q(d\nu) p(a, dw) V_{n-1}(x - y + w, w, \nu)] \tag{61}$$

Let $\mathbf{v_1} = (x + \Delta x, a + \Delta a, \mu)$, $\mathbf{v_2} = (x, a, \mu)$. Let the minimum in (61) corresponding to $\mathbf{v_1}$, $\mathbf{v_2}$ be attained, respectively, at y_1, y_2.

Case 1: $y_1 - y_2 \leq \Delta x$.

$$y_1 - y_2 \leq \Delta x$$

$$\Rightarrow \quad x + \Delta x - y_1 + w \geq x - y_2 + w$$

$$\Rightarrow^{*^1} \quad [V_{n-1}(x + \Delta x - y_1 + w, w, \mu) - V_{n-1}(x - y_2 + w, w, \mu)] \uparrow \text{ as } w \uparrow$$

$$\Rightarrow^{*^2} \quad \int q(d\nu)p(a + \Delta a, dw)[V_{n-1}(x + \Delta x - y_1 + w, w, \nu) - V_{n-1}(x - y_2 + w, w, \nu)]$$

$$\geq \int q(d\nu)p(a, dw)[V_{n-1}(x + \Delta x - y_1 + w, w, \nu) - V_{n-1}(x - y_2 + w, w, \nu)]$$

$$\Rightarrow \quad V_n(x + \Delta x, a + \Delta a, \mu) + V_n(x, a, \mu)$$

$$= \lambda(x + \Delta x) + \mu F(y_1) + \beta \int q(d\nu)p(a + \Delta a, dw)V_{n-1}(x + \Delta x - y_1 + w, w, \nu)$$

$$+ \lambda x + \mu F(y_2) + \beta \int q(d\nu)p(a, dw)V_{n-1}(x - y_2 + w, w, \nu)$$

$$\geq \lambda x + \mu F(y_2) + \beta \int q(d\nu)p(a + \Delta a, dw)V_{n-1}(x - y_2 + w, w, \nu)$$

$$+ \lambda(x + \Delta x) + \mu F(y_1) + \beta \int q(d\nu)p(a, dw)V_{n-1}(x + \Delta x - y_1 + w, w, \nu)$$

$$= h_n(y_2, x, a + \Delta a, \mu) + h_n(y_1, x + \Delta x, a, \mu)$$

$$\geq V_n(x, a + \Delta a, \mu) + V_n(x + \Delta x, a, \mu)$$

where $(*^1)$ follows by Theorem 1, the induction hypothesis that (10) holds for V_{n-1}, and Lemma 14. $(*^2)$ follows by Lemma 1 and our stochastic monotonicity assumption on $a \to p(a, \cdot)$.

Case 2: $y_1 - y_2 \geq \Delta x$.

$$y_1 - y_2 \geq \Delta x$$

$$\Rightarrow^* F(y_1) + F(y_2) \geq F(y_1 - \Delta x) + F(y_2 + \Delta x)$$

$$\Rightarrow \quad V_n(x + \Delta x, a + \Delta a, \mu) + V_n(x, a, \mu)$$

$$= \lambda(x + \Delta x) + \mu F(y_1) + \beta \int q(d\nu)p(a + \Delta a, dw)V_{n-1}(x + \Delta x - y_1 + w, w, \nu)$$

$$+ \lambda x + \mu F(y_2) + \beta \int q(d\nu)p(a, dw)V_{n-1}(x - y_2 + w, w, \nu)$$

$$\geq \lambda x + \mu F(y_1 - \Delta x) + \beta \int q(d\nu)p(a + \Delta a, dw)V_{n-1}(x + \Delta x - y_1 + w, w, \nu)$$

$$+ \lambda(x + \Delta x) + \mu F(y_2 + \Delta x) + \beta \int q(d\nu)p(a, dw)V_{n-1}(x - y_2 + w, w, \nu)$$

$$= h_n(y_1 - \Delta x, x, a + \Delta a, \mu) + h_n(y_2 + \Delta x, x + \Delta x, a, \mu) \geq V_n(x, a + \Delta a, \mu) + V_n(x + \Delta x, a, \mu)$$

(*) follows since F is convex. Thus, V_0 satisfies the assertion of the theorem, and V_n satisfies the assertion of the theorem on the hypothesis that V_{n-1} does. By induction, the theorem is proved. ∎

REFERENCES

[1] B. Ata, "Dynamic power control in a wireless static channel subject to a quality of service constraint," *Operations Research*, vol. 53, pp. 842–851, 2005.

[2] R. Berry, *Power and Delay Trade-offs in Fading Channels*. PhD thesis, Laboratory for Information and Decision Systems, M.I.T, 2000.

[3] R. Berry and R. Gallager, "Communication over fading channels with delay constraints," *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1135–1149, 2002.

[4] B. Collins and R. Cruz, "Transmission policies for time varying channels with average delay constraints," in *Proc. 37th Allerton Conf. on Communication, Control and Computing, Monticello, IL*, 1999.

[5] M. Goyal, A. Kumar, and V. Sharma, "Power constrained and delay optimal policies for scheduling transmission over a fading channel," in *Proc. INFOCOM*, 2003.

[6] G. Rajadhyaksha and V. Borkar, "Transmission rate control over randomly varying channels," *Prob. in the Engg. and Info. Sci.*, vol. 19, pp. 73–82, 2005.

[7] D. Rajan, A. Sabharwal, and B. Aazhang, "Delay-bounded packet scheduling of bursty traffic over wireless channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 125–144, 2004.

[8] E. Uysal-Biyikoglu and A. E. Gamal, "On adaptive transmission for energy efficiency in wireless data networks," *IEEE Transactions on Information Theory*, vol. 50, pp. 3081–3094, 2004.

[9] E. Uysal-Biyikoglu, B. Prabhakar, and A. E. Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Transactions on Networking*, vol. 10, pp. 487–499, 2002.

[10] H. Wang, *Opportunistic Transmission of Wireless Data over Fading Channels under Energy and Delay Constraints*. PhD thesis, Rutgers University, 2003.

[11] D. V. Djonin and V. Krishnamurthy, "Structural results on the optimal transmission scheduling policies and costs for correlated sources and channels," in *IEEE CDC*, 2005.

[12] S. P. Meyn, "The policy improvement algorithm for Markov decision processes with general state space," *IEEE Transactions on Automatic Control*, vol. 42, pp. 191–196, 1997.

[13] E. Altman and G. Koole, "On submodular value functions and complex dynamic programming," *Stochastic Models*, vol. 14, pp. 1051–1072, 1998.

[14] G. Koole, "Structural results for the control of queueing systems using event-based dynamic programming," *Queueing Systems*, vol. 30, pp. 323–339, 1998.

[15] J. G. Shanthikumar and D. Yao, "Multiclass queueing systems: polymatroidal structures and optimal scheduling control," *Operations Research*, vol. 4, pp. S293–S299, 1992.

[16] J. Smith and K. McCardle, "Structural properties of stochastic dynamic programs," *Operations Research*, vol. 50, pp. 796–809, 2002.

[17] O. Hernández-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes*. Springer Verlag, New York, 1996.

[18] O. Hernández-Lerma and J. B. Lasserre, *Further Topics on Discrete-Time Markov Control Processes*. Springer Verlag, New York, 1999.

[19] C. C. Pugh, *Real Mathematical Analysis*. Springer Verlag, New York, 2002.

[20] R. Gallager and D. Bertsekas. Prentice Hall, 1991.

[21] M. Shaked and J. G. Shanthikumar, *Stochastic Orders and their Applications*. Academic Press, New York, 1994.

[22] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*. Springer Verlag, New York, 1993.

[23] V. Borkar, "Convex analytic methods in Markov decision processes," in *Handbook of Markov Decision Processes* (E. Feinberg and A. Shwartz, eds.), pp. 347–375, Kluwer Academic Publ., Norwell, MA.

[24] E. Altman, *Constrained Markov Decision Processes*. Chapman and Hall/CRC, Boca Raton, Fl, 1999.

[25] O. Hernández-Lerma and J. B. Lasserre, "Approximation schemes for infinite linear programs," *SIAM J. Optim.*, vol. 8, pp. 973–988, 1998.

[26] D. P. de Farias and B. V. Roy, "The Linear programming approach to approximate Dynamic Programming," *Operations Research*, vol. 51, no. 6, pp. 850–865, 2003.

[27] V. Borkar, "An actor-critic algorithm for constrained Markov Decision Process," *Systems and Control Letters*, vol. 54, pp. 207–213, 2005.

[28] H. L. Royden, *Real Analysis*. Prentice Hall, Englewood Cliffs, NJ, 4th ed., 2005.

[29] P. Glynn and D. Ormoneit, "Hoeffding's inequality for uniformly ergodic Markov chains," *Statistics and Probability Letters*, vol. 56, no. 2, pp. 143–146, 2002.

[30] R. Sznajder and J. Filar, "Some comments on a theorem of Hardy and Littlewood," *J. Opt. Theory and Appl.*, vol. 75, pp. 201–208, 1992.

[31] R. Sundaram, *A First Course in Optimization Theory*. Cambridge University Press, Cambridge, UK, 1996.